10/565163
IAP12 Rec'd PCT/PTO 19 JAN 2006

AN AUDIO STEREO PROCESSING METHOD, DEVICE AND SYSTEM

FIELD OF THE INVENTION

The present invention relates to a method, a device and a system for processing an audio stereo signal, and in particular the present invention relates to a method, device and system for processing an input audio stereo signal in accordance with the preamble of claims 1, 14 and 27, respectively.

BACKGROUND OF THE INVENTION

A large number of methods and systems exist intended for faithful reproduction of the sound experienced by a listener at the recording position. The system coming closest to virtually move the listener to the recording location, i.e. to convey an impression of the true location of the different sound sources of the original event, is the binaural method of recording and the binaural method of reproduction (headphones). This method has as its shortcomings in that the sound is interpreted by ear canals both in the recording stage and in the reproduction stage and in a worst case even by two sets of pinna (outer ears) on its way to the listeners brain where the sound information is to be interpreted. There are solutions that utilize a simplified recording method including a foam ball of head size with the microphone elements on each side of the ball instead of a replica of a head. This is a compromise to gain sound quality but loses the distinction of localization between front and back and the elevation. All other ways than the binaural method to record and reproduce sound is a creation of an imaginary sound image that is truly subjective. This is the case for both the recording stage and the reproduction stage.

As opposed to previously known methods, the object of the reproduction stage should only be to convey the electrical differences to the listener's auditory system with minimal loss or addition of information. The place where the stereo sound image is created is then the recording and/or mixing stage. The stereo image might be made as a truthful, but still subjective, interpretation of the sound experienced by a listener in a venue, or as an illusion of an imaginary event that never have physically occurred or a mixture of the two.

Most reproduction systems of today are based on a pair of widely spaced loudspeakers, and true reproduction of the electrical stereo signal, both in terms of relative intensity between the sound waves perceived by the ears of the listener and the time difference between these, can at best be perceived only at one single position in relation to the loudspeakers. These methods are often subject to incorrect translation of the electrical stereo information dependant on the preferences of the separate loudspeakers and how the loudspeakers are positioned in relation to the listener. There is thus a need for a sound reproduction system that provides identical reproduction of the stereo sound image regardless of setup and quality of the loudspeakers.

A system that solves this problem is described in the patent application WO01/39548, assigned to the applicant of the present invention, which discloses a method of processing and reproducing an input audio stereo signal. A side signal is split into a first and a second intermediate signal, where the first intermediate signal is equal to the side signal and the second intermediate signal is equal to the first intermediate signal phase shifted 180°, a mid signal is attenuated by a factor α which compensates for imperfections in the balance

between the mid and side signals appearing in the audio reproduction stage, the attenuated mid signal is added to both the first and the second intermediate signals, so as to form the output audio stereo signal, and the output stereo signal is directed to an audio stereo signal reproduction system comprising a pair of loudspeaker units located in close proximity to each other. The system described in WO01/39548 allows an audio stereo signal to be reproduced with a high degree of fidelity with high consistency in the perceived stereo image regardless of the quality of system.

A problem with such a system with closely located loudspeaker units, however, is that at high frequencies, above 1-5 kHz, the degree of fidelity in perceived stereo effect degrades or vanishes totally.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a method for processing an audio stereo signal, which solves the above mentioned problem. This object is achieved by a method as defined in the characterising portion of claim 1.

Another object of the present invention is to provide a device for processing an audio stereo signal, which solves the above mentioned problem. This object is achieved by a device as defined in the characterising portion of claim 14.

Another object of the present invention is to provide a system for processing an audio stereo signal, which solves the above mentioned problem. This object is achieved by a system as defined in the characterising portion of claim 27.

According to the present invention, a left output signal for transmission to a left loudspeaker in a loudspeaker pair is produced, which signal is, or is equivalent to, the sum of a

mid input signal M and a side input signal S, at least part of which side signal S or mid signal M being phase shifted approximately 45°-135° with respect to the other signal, and a right output signal for transmission to a right loudspeaker in said pair is produced, which signal is, or is equivalent to, the sum of the mid input signal M, and a 180° phase shifted side signal S, at least part of which side signal S or mid signal M being phase shifted approximately 45°-135° with respect to the other signal.

This has the advantage that the phase difference that the present invention introduces into the stereo signal translates incoming level difference into phase difference between the stereo channels. This phase difference will be translated into a level difference when the stereo signal is played back through a loudspeaker pair. Level difference, in contrast to phase difference, is a strong localization cue for shorter wavelengths, and consequently the phase shift introduced by the present invention will improve the degree of fidelity in perceived stereo effect considerably.

The mid input signal M may be attenuated by a factor α and/or the side input signal S may be amplified a factor β in the production of the left output signal and the right output signal. This has the advantage that a stereo audio signal composed of level difference for long wavelengths and phase difference for short wavelengths may be obtained, which signal will be played back through a loudspeaker pair as phase difference for low frequencies, which is a strong localization cue for low frequencies, and level difference for high frequencies, which, as mentioned above, is a strong localization cue for high frequencies.

The input signals in the present invention may be a left input signal L and a right input signal R, in which case the mid input signal M is produced as the sum of the left input signal L and the right input signal R, and the side input signal is produced as the difference of the left input signal L and the right input signal R. This has the advantage that a conventional stereo signal may be used as input signals in the present invention.

The loudspeaker elements may be closely located, and in particular the pair of loudspeaker elements may consist of a pair of identical loudspeaker elements being acoustically isolated from each other, and located within less than one quarter of the shortest wavelength emitted by the elements, or, if the shortest wavelength emitted by the elements is less than 68 cm, less than 17 cm. This has the advantage that the present invention is very well suited for use in a method and system as described in WO01/39548.

The phase shift may be accomplished such that all of the side input signal S or the mid input signal M is phase shifted 45°-135°, preferably 90°. This may advantageously be accomplished by digital signal processing, e.g. by a Hilbert transform. Alternatively, the phase shift may be accomplished by a frequency dependent filter, such as an analogue all-pass filter. This has the advantage that a less expensive solution may be obtained for cost sensitive applications and/or applications where the processing time is critical.
The mid input signal M may be delayed a time corresponding to the delay of the phase shifting means. This may facilitate the obtaining of a desired phase relation between the side input signal S and the mid input signal M.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram illustrating a prior art system for processing stereo signals;

Fig. 2 is a block diagram illustrating a first embodiment of the present invention;

Fig. 3 is a block diagram illustrating a second embodiment of the present invention; and

Fig. 4 shows an example of the frequency response of an all pass filter in the embodiment shown in fig. 3.

DESCRIPTION OF A PREFERRED EMBODIMENT

Fig. 1 illustrates the functional principle for a prior art device for processing an audio stereo signal. The input audio stereo signal comprises a left input stereo signal L and a right input stereo signal R. The L and R signals are used to obtain a mid signal M, and a side signal S, corresponding to the sum of the left L and right R input stereo signals, and the difference between the left L and right R input stereo signals, respectively. The output stereo signal L_{OUT}, which is to be sent to a left sound reproducing unit, such as a loudspeaker, is the sum of the side signal, S, and the mid signal M multiplied by an attenuating factor α, while the output stereo signal R_{OUT}, which is to be sent to a right sound reproducing unit is the sum of the inverted side signal, S, and the mid signal M multiplied by an attenuating factor α.

The system described in fig. 1 allows an electrical audio stereo signal to be reproduced with a high degree of fidelity with high consistency in the perceived stereo image regardless of the quality of system. As stated above, however, the system

in fig. 1 suffers from the problem that the degree of fidelity in perceived stereo effect degrades or vanishes totally at frequencies above 1-5 kHz.

This is due to the fact that level difference in L_{OUT} and R_{OUT} resulting from the respective addition and subtraction of the S signal is transformed into phase difference when played back through the loudspeaker elements. This phase difference is a strong localization cue for low frequencies, and results in excellent stereo resolution for these lower frequencies. Due to the characteristics of the human ear, however, the ability to detect phase differences between two signals received by the left and the right ear, respectively, vanishes at high frequencies. The reason for this is the phase locking of the auditory nerve that tend to fire at a particular phase of a stimulating low frequency tone (< 4 -5 kHz), with one burst of spikes per cycle for frequencies below about 1000 Hz. Inter-spike intervals tend to occur at integer multiples of the period of the tone. With high frequency tones (> 4 - 5 kHz) phase locking gets weaker and then disappear, because the capacitance of inner hair cells prevents them from changing in voltage sufficiently rapidly. The lack of phase locking above 4-5 kHz affirms that the system in fig. 1 conveys week localization cues for sounds containing only short wavelengths with only level difference between the stereo channels.

The present invention seeks to solve the above problem with a device as illustrated in fig. 2. The device in fig. 2 is similar to the device in fig. 1 with the difference that in fig. 2 an extra unit 20 has been added. As in fig. 1, a mid signal M is obtained by summing the left L and right R input stereo signals, and a side signal S is obtained by subtracting the right input stereo signal R from the left input stereo

signal L. The side signal S is then phase shifted -90° prior to the creation of the output stereo signals L_{OUT}, and R_{OUT}. The output stereo signal L_{OUT} is then obtained by taking the sum of the phase shifted side signal S and the mid signal M multiplied by an attenuating factor α, while the output stereo signal R_{OUT} is obtained by subtracting the phase shifted side signal S from the mid signal M multiplied by an attenuating factor α. This is equal to taking the sum of an inverted phase shifted side signal S and the mid signal M multiplied by an attenuating factor α. Inverting the side signal is equivalent to negating it or phase shifting it 180°.

The attenuation factor α would typically be -6 dB to -12dB. In a general case, however, the attenuation factor α is adapted to optimise the stereo effect perceived by the listener, and is allowed to vary in an interval from -3 dB to -15 dB.

The phase shift may be accomplished by a digital signal processor, e.g. by a Hilbert transform. Digital signal processing has the advantage that a true 90° phase shift can be performed for all wavelengths and may be obtained with little or no amplitude change over frequency (use of analogue circuits may result in a phase drift in the audible spectra in the range of 500°-700° or more, however with a relative phase difference of 90° between the mid signal M and the side signal S). This type of phase shifting is particularly suitable for systems in which digital signal processing means already are present, and where the applications are not time critical.

Further, it may be desirable to include a delay circuit in the device, shown as 21 in fig. 2, to delay the mid input signal M with a time corresponding to the processing time of the phase shifting means. This facilitates maintaining of the desired

phase relation between the side input signal S and the mid input signal M.

Fig. 3 illustrates a second embodiment of the present invention. The second embodiment of the present invention is a solution for applications where the phase shift is desired but the application is cost sensitive and/or where the processing time is critical, such as in professional recording studios. In the second embodiment, the mid signal M and side signal S is obtained as in fig. 2, and the side signal is then S altered by a unit 30 including a frequency dependent analogue all pass filter with its centre frequency set well above the shortest audible wavelength. This means that the phase shift is starting with only a few degrees at e.g. 500 Hz to reach +90° at e.g. 10 kHz. The phase response of the all pass filter is thus tailored to gradually translate the phase difference of the output stereo signal into level difference as the phase locking gets weaker for higher frequencies. As the phase response of an analogue filter rarely can be of negative nature the unit 30 further includes means to invert the signal to get the desired result of a phase shift of -90°. The phase shift should preferably be negative since otherwise the original L and R signals might be switched. An example of a phase response for the all pass filter is shown in fig. 4. As can be seen in the figure, the phase shift starts from substantially 0° at low frequencies to reach 90° at high frequencies (e.g. 10 kHz). It is also possible to create the frequency dependent phase shift with the aid of digital signal processing, however with the extra cost this might incur.

The factor α in fig. 3 can be made frequency dependant so that the factor is different for separate drivers of, for example, different elements in a multi-way loudspeaker configuration.

The mid signal M is then added to the phase shifted side signal S to form a first output signal, and the phase shifted side signal S is then subtracted from the mid signal M to form the second output signal.

Generally, the method described in the present application could equivalently be used for any input terms which can be described as a linear transformation of the R and L signals or the M and S signals, but as a matter of convenience, the method has been exemplified using the M and S, and the R and L pictures, respectively. The method should therefore be interpreted as a method having an output, which is equivalent to $S_{ps}+\alpha M$ and $-S_{ps}+\alpha M$, where S_{ps} is the S signal phase shifted with 90°. As has been described, the M and S signals may be produced during an intermediate step in the process, but this does not have to be the case as long as the resulting output condition is fulfilled.

In the above description the phase shift has been described as 90°. This phase shift may however be any phase shift in an interval between 45°-135°. Further, in the above description the phase shift has been performed on the side signal S. It may however as well be performed on the mid signal M.

Further, in the above description the analogue all pass filter could however be exchanged by a digital filter doing an identical filtering function as the above described analogue all pass filter. In this case, it may be desirable to include a delay circuit in the device, as shown as 21 in fig. 2, to delay the mid input signal M with a time corresponding to the processing time of the phase shifting means also in this embodiment.

Further, in the above description the input stereo signals consist of a L and a R signal. The input signals could however as well consist of the M and S signals, in which case the first addition and subtraction steps are omitted.

Further, in the above description the mid signal M has been attenuated a factor α. It is, however, of course possible to amplify the side signal S with a factor β instead.

In the detailed description of the present invention the phase shift has been carried out on the side input signal S. The phase shift could however as well be carried out on the mid input signal M.

Inasmuch as the present invention is subject to variations, modifications and changes in detail, some of which have been stated herein, it is intended that all matter described throughout this entire specification or shown in the accompanying drawings be interpreted as illustrative and not in a limiting sense.